LISTING OF THE CLAIMS

Claims 1-30 are pending. No claims are amended. No claims are cancelled or added. The following listing of claims replaces all prior versions and listings of claims in the application.

1. (Original) A method for clustering queries, the method comprising:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries; and

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that the queries are similar independent of whether individual ones of the queries comprise similar composition with respect to other ones of the queries.

2. (Original) A method as recited in claim 1, wherein the queries comprise a well formed natural language question, a keyword, or a phrase.

3. (Original) A method as recited in claim 1, wherein the query cluster is used to disambiguate a word or phrase in a query of the queries.

4. (Original) A method as recited in claim 1, further comprising determining that the queries are similar based on similar keyword or phrase composition.

5. (Original) A method as recited in claim 1, wherein identifying the same document and/or the similar documents further comprises:

determining the similar documents by evaluating a set of selected similar documents chosen responsive to queries p and q of the queries, wherein documents $D\_C(.)$ is a subset of a result list $D(.)$ according to the following:

$$D\_C(p) = \{ d_{p1}, d_{p2}, \ldots, d_{pi} \} \subseteq D(p)$$

$$D\_C(q) = \{ d_{q1}, d_{q2}, \ldots, d_{qj} \} \subseteq D(q);$$

wherein similarity based on selection of documents is based on:

If $D\_C(p) \cap D\_C(q) = \{ d_{pq1}, d_{pq2}, \ldots, d_{pqk} \} \neq \varnothing$, then documents d_{pq1}, d_{pq2}, ..., d_{pqk} represent a set of common topics of queries p and q, and,

whereby the similar documents between queries p and q is determined by $D\_C(p) \cap D\_C(q)$.

6. (Original) A method as recited in claim 1, further comprising constructing a thesaurus comprising a plurality of synsets, wherein each synset comprises one or more query clusters.

7. (Original) A method as recited in claim 1, wherein identifying the same document and/or the similar documents further comprises determining the similar documents based on a proportionality of commonly selected individual documents.

8. (Original) A method as recited in claim 7, wherein identifying the same document and/or the similar documents further comprises:

determining the similar documents based on a proportionality of commonly selected individual documents, such that:

$$similarity_{single\_doc}(p, q) = \frac{RD(p, q)}{Max(rd(p), rd(q))},$$

wherein $rd(.)$ is the number of clicked documents for a query of the queries, and wherein $RD(p, q)$ is the number of document selections in common.

9. (Original) A method as recited in claim 1, wherein identifying the same document and/or the similar documents further comprises:

determining the similar documents based on a hierarchical positioning between individual ones of a plurality of documents commonly selected across the queries.

10. (Original) A method as recited in claim 9:

wherein $F(d_i, d_j)$ is a lowest common parent node for documents d_i and d_j;

wherein $L(x)$ is a level of a node x;

wherein L_Total identifies a total number of levels in a hierarchy; and

wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that}$$

$s(d_i, d_i) = 1$; and $s(d_i, d_j) = 0$ if $F(d_i, d_j) =$ root; and

the method further comprises:

incorporating $s(d_i, d_j)$ into a calculation of query similarity, wherein. d_i $(1 \le i \le m)$ and d_j $(1 \le j \le n)$ be a set of selected documents for queries p and q respectively such that:

$$similarity_{hierarchy}(p, q) = \frac{1}{2} \times \left(\frac{\sum_{i=1}^{m} (\max_{j=1}^{n} s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^{n} (\max_{i=1}^{m} s(d_i, d_j))}{rd(q)} \right)$$

11. (Original) Computer-readable media comprising computer-executable instructions for identifying similar queries, the computer-executable instructions comprising instructions for:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries; and

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that the queries are similar independent of

whether individual ones of the queries comprise similar composition with respect to other ones of the queries.

12. (Original) Computer-readable media as recited in claim 11, wherein the queries comprise a well formed natural language question, a keyword, or a phrase.

13. (Original) Computer-readable media as recited in claim 11, wherein the query cluster is used to disambiguate a word or phrase in a query of the queries.

14. (Original) Computer-readable media as recited in claim 11, wherein the computer-executable instructions further comprise instructions for determining that the queries are similar based on similar keyword or phrase composition.

15. (Original) Computer-readable media as recited in claim 11, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:

determining the similar documents by evaluating a set of selected similar documents chosen responsive to queries p and q of the queries, wherein documents $D\_C(.)$ is a subset of a result list $D(.)$ according to the following:

$$D\_C(p) = \{ d_{p1}, d_{p2}, \ldots, d_{pi} \} \subseteq D(p)$$

$$D\_C(q) = \{ d_{q1}, d_{q2}, \ldots, d_{qj} \} \subseteq D(q);$$

wherein similarity based on selection of documents is based on:

If $D\_C(p) \cap D\_C(q) = \{ d_{pq1}, d_{pq2}, \ldots, d_{pqk} \} \neq \varnothing$, then documents d_{pq1}, d_{pq2}, ..., d_{pqk} represent a set of common topics of queries p and q, and,

whereby the similar documents between queries p and q is determined by $D\_C(p) \cap D\_C(q)$.

16. (Original) Computer-readable media as recited in claim 11, wherein the computer-executable instructions further comprise instructions for constructing a thesaurus comprising a plurality of synsets, wherein each synset comprises one or more query clusters.

17. (Original) Computer-readable media as recited in claim 11, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for determining the similar documents based on a proportionality of commonly selected individual documents.

18. (Original) Computer-readable media as recited in claim 17, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:

determining the similar documents based on a proportionality of commonly selected individual documents, such that:

$$similarity_{single\_doc}(p, q) = \frac{RD(p, q)}{Max(rd(p), rd(q))},$$

wherein $rd(.)$ is the number of clicked documents for a query of the queries, and wherein $RD(p, q)$ is the number of document selections in common.

19. (Original) Computer-readable media as recited in claim 11, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:

determining the similar documents based on a hierarchical positioning between individual ones of a plurality of documents commonly selected across the queries.

20. (Original) Computer-readable media as recited in claim 19:

wherein $F(d_i, d_j)$ is a lowest common parent node for documents d_i and d_j;

wherein $L(x)$ is a level of a node x;

wherein L_Total identifies a total number of levels in a hierarchy; and

wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that}$$

$s(d_i, d_i) = 1$; and $s(d_i, d_j) = 0$ if $F(d_i, d_j) =$ root; and

wherein the computer-executable instructions further comprise instructions for:

incorporating $s(d_i, d_j)$ into a calculation of query similarity, wherein. d_i $(1 \le i \le m)$ and d_j $(1 \le j \le n)$ be a set of selected documents for queries p and q respectively such that:

$$similarity_{hierarchy}(p, q) = \frac{1}{2} \times \left(\frac{\sum_{i=1}^{m} (\max_{j=1}^{n} s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^{n} (\max_{i=1}^{m} s(d_i, d_j))}{rd(q)} \right)$$

21. (Original) A computing device comprising:

a processor coupled to a memory, the memory comprising computer executable instructions, the processor being configured to fetch and execute the computer-executable instructions for:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries; and

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that the queries are similar independent of whether individual ones of the queries comprise similar composition with respect to other ones of the queries.

22. (Original) A computing device as recited in claim 21, wherein the queries comprise a well formed natural language question, a keyword, or a phrase.

23. (Original) A computing device as recited in claim 21, wherein the query cluster is used to disambiguate a word or phrase in a query of the queries.

24. (Original) A computing device as recited in claim 21, wherein the computer-executable instructions further comprise instructions for determining that the queries are similar based on similar keyword or phrase composition.

25. (Original) A computing device as recited in claim 21, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:

determining the similar documents by evaluating a set of selected similar documents chosen responsive to queries p and q of the queries, wherein documents $D\_C(.)$ is a subset of a result list $D(.)$ according to the following:

$$D\_C(p) = \{ d_{p1}, d_{p2}, \ldots, d_{pi} \} \subseteq D(p)$$

$$D\_C(q) = \{ d_{q1}, d_{q2}, \ldots, d_{qj} \} \subseteq D(q);$$

wherein similarity based on selection of documents is based on:

If $D\_C(p) \cap D\_C(q) = \{ d_{pq1}, d_{pq2}, \ldots, d_{pqk} \} \neq \varnothing$, then documents d_{pq1}, d_{pq2}, ..., d_{pqk} represent a set of common topics of queries p and q, and,

whereby the similar documents between queries p and q is determined by $D\_C(p) \cap D\_C(q)$.

26. (Original) A computing device as recited in claim 21, wherein the computer-executable instructions further comprise instructions for constructing a thesaurus comprising a plurality of synsets, wherein each synset comprises one or more query clusters.

27. (Original) A computing device as recited in claim 21, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for determining the similar documents based on a proportionality of commonly selected individual documents.

28. (Original) A computing device as recited in claim 27, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:

determining the similar documents based on a proportionality of commonly selected individual documents, such that:

$$similarity_{single\_doc}(p, q) = \frac{RD(p, q)}{Max(rd(p), rd(q))},$$

wherein $rd(.)$ is the number of clicked documents for a query of the queries, and wherein $RD(p, q)$ is the number of document selections in common.

29. (Original) A computing device as recited in claim 21, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:

determining the similar documents based on a hierarchical positioning between individual ones of a plurality of documents commonly selected across the queries.

30. (Original) A computing device as recited in claim 29:

wherein $F(d_i, d_j)$ is a lowest common parent node for documents d_i and d_j;

wherein $L(x)$ is a level of a node x;

wherein L_Total identifies a total number of levels in a hierarchy; and

wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that}$$

$s(d_i, d_i) = 1$; and $s(d_i, d_j) = 0$ if $F(d_i, d_j) =$ root; and

wherein the computer-executable instructions further comprise instructions for:

incorporating $s(d_i, d_j)$ into a calculation of query similarity, wherein. d_i $(1 \le i \le m)$ and d_j $(1 \le j \le n)$ be a set of selected documents for queries p and q respectively such that:

$$similarity_{hierarchy}(p, q) = \frac{1}{2} \times \left(\frac{\sum_{j=1}^{m} \left(\max_{j=1}^{n} s(d_i, d_j) \right)}{rd(p)} + \frac{\sum_{j=1}^{n} \left(\max_{i=1}^{m} s(d_i, d_j) \right)}{rd(q)} \right)$$